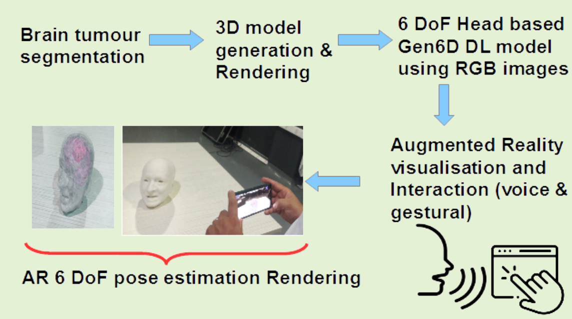
HOLOTumour: 6DoF Phantom Head pose estimation based deep learning for brain tumour segmentation and AR visualisation and interaction

K. Amara, M.A. Guerroudji, O. Kerdjidj, N. Zenati, and N. Ramzan



***Abstract*—To comprehend a scene, and use and interact with objects in an augmented reality (AR) application correctly, 6 DoF (Degrees of Freedom) object pose estimation is a crucial task. By supplying exact 3D pixel coordinates, the incorporation of depth photos has resulted in substantial gains. However, depth photos are not always simple to find; for instance, mobile phones and tablets, which are common devices for augmented reality apps, do not provide any depth information. As a result, a lot of study is focused on determining the poses of known objects using simple RGB images. 3D Poorly textured objects detection is a challenging task due to their lack of distinctive features, limited appearance variations, ambiguity with the background, occlusion challenges, lighting and reflection effects, and lack of distinct and unique features. In this study, we created a framework to automate the workflow for predicting brain tumour segmentation and 6 DoF localisation. Our framework named 'HOLOTumour' provides an augmented-reality rendering with 3D AR visualisation and interaction, generates the 3D model of the brain with a segmented tumour based on Geodesic-Aided Chan-Vese Model employing a local Magnetic Resonance Imaging (MRI) dataset, and outputs the phantom head 6 DoF pose estimation using generalisable object pose estimators employing only RGB images. Based on the comparison study including the accuracy and inference runtimes scores, the 6 DoF pose estimation of the printed phantom head and brain tumour segmentation achieve promising performance. Medical professionals may identify brain tumours with the help of the HOLOTumour platform, which can also be used for professional practice and training in medicine.**

***Index Terms*—6DoF pose estimation, Deep Learning, RGB images, Gen6D, brain tumour segmentation, MRI, augmented reality, 3D reconstruction, Phantom head, 3D visualisation, voice interaction, gestural Interaction.**

I. Introduction

OBJECT pose estimation determines the position and orientation of a 3D object in the real world, based on its appearance in images or video frames. The goal of object pose estimation is to accurately localise the object in 3D space relative to a known coordinate system, such as the camera or the environment, using various computer vision techniques and algorithms. 6 DoF pose estimation permits machines to perceive their surroundings in 3D, allowing them to accomplish tasks such as object tracking, recognition, and manipulation. For many tasks involving interaction with objects, estimating the orientation and placement of the object in 3D space is a preliminary and essential stage. Robotics, video games, and immersive technologies like augmented and virtual realities have all benefited greatly from the remarkable advancements in 3D vision over the past last years. The 6 DoF object pose estimation is put to additional demands by these applications. The aforementioned applications place new demands on the 6 DoF object pose estimation, necessitating the development of a pose estimator that is flexible, real-time, generalizable, and simple to use.

A significant amount of study has been devoted to the topic of 3D object detection since it has always been crucial to computer vision. With the development of deep learning, this issue, along with many others related to vision, underwent a complete revival. In more recent efforts, accuracy and efficiency were equally important in order to make approaches adaptable to real-world conditions with computational and runtime restraints.

Traditionally, feature points from images and 3D models are matched to solve the 6 DoF object pose estimation problem.

K. Amara, M.A. Guerroudji, O. Kerdjidj, and N. Zenati, are with CDTA-Centre for Development of Advanced Technologies, Algiers, Algeria (kamara@cdta.dz). O. Kerdjidj is with the College of Engineering and Information Technology, University of Dubai, UAE. N. Ramzan is with the School of Engineering and Computing, University of the West of Scotland, Paisley, Scotland, United Kingdom
.

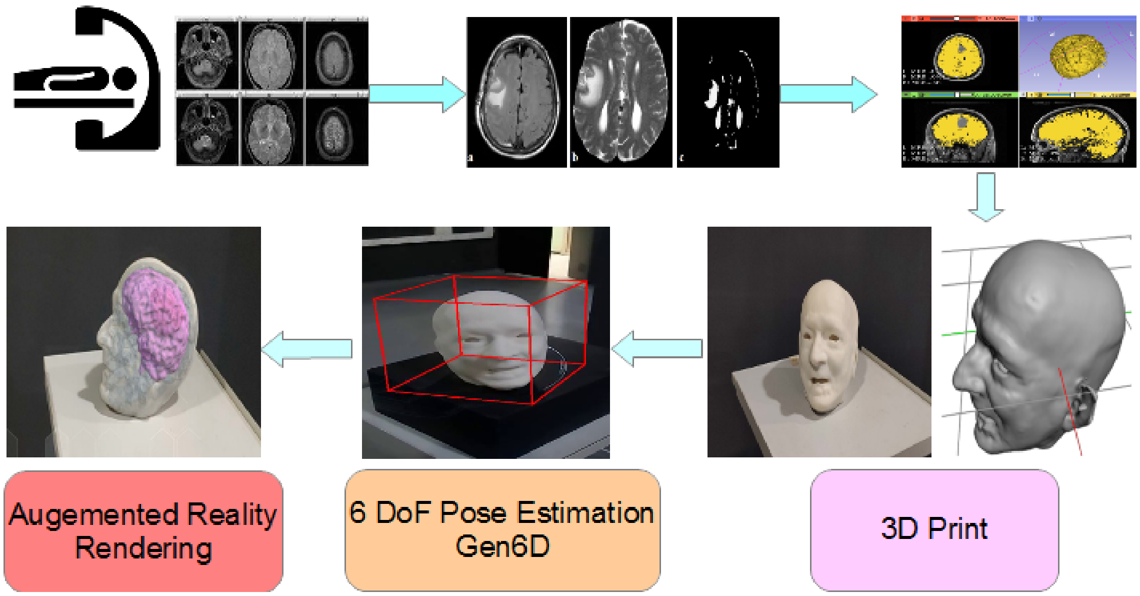


Fig. 1: HOLOTumour general architecture for AR medical practice

However, to identify feature points for matching, these techniques require a rich texture of the objects [1]. Consequently, these methods cannot manage texture-less or poorly textured objects. With the advent of depth cameras, several techniques for 3D texture-less objects detection using RGB-D data have been proposed.

In the last two years, 6 DoF object pose estimation using RGB images-based deep learning approaches have shown impressive performance. There are several architectures for deep learning-based methods, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and hybrid models. One advantage of these methods is their ability to estimate the pose of objects in real time, making them suitable for applications such as robotics, AR, and autonomous driving. In this section, we will highlight the latest 3D pose estimation methods. We made sure to only pick methods that are computer-aided design CAD-model-free, meaning they don't implement rendering the 3D model during the inference phase, and only require RGB images as input. Here we review the following ones: OnePose++ [2], PvNet [3], EfficientPose [4], and MediaPipe Objectron [5]:

- OnePose++ [2] is based on the keypoint-matching-based method called "OnePose" [6], which has shown promising results under a one-shot setting, where it relies on detecting repeatable image keypoints, which eliminates the need for CAD models but is thus prone to failure on low-textured objects. In OnePose++, a keypoint-free pose estimation pipeline is implemented to remove the need for repeatable keypoint detection. Given a query image for object pose estimation, a 2D-3D matching network directly establishes 2D-3D correspondences between the query image and the reconstructed point-cloud model without first detecting key points in the image. Experiments show that the OnePose++ outperforms existing one-shot CAD-model-free methods in precision by a wide margin, and is comparable to CAD-model-based methods on the LINEMOD dataset [7].
- PVNet [3] (Pixel-wise Voting Network) is a 6D object pose estimation method. It regresses pixel-wise unit vectors pointing to the key points and uses these vectors to vote for key point locations using RANSAC.
- EfficientPose [4] is a 3D object pose estimation method. It can detect the 2D bounding box of multiple objects and instances as well as estimate their full 3D pose from a single RGB image. First, it detects 2D targets, via key points, then it solves a Perspective-n-Point problem. This method was extended from the 2D object detection method EfficientDet to also predict the 3D pose of objects, which is the translation and rotation of the object in the image. EfficientPose network consists of: the EfficientNet backbone, the bidirectional feature pyramid network (BiFPN) and the prediction subnetworks.
- MediaPipe Objectron [5] is a closed-source pre-built module in the MediaPipe framework, it's a 3D object pose estimation method from RGB images input. It uses machine learning models to estimate the 3D bounding box of objects in a video stream or a sequence of images. It's designed and trained to detect objects with a predefined set of 4 classes: shoes, chairs, cups, and cameras. Objectron uses a combination of 2D image features and geometric priors to estimate the 3D pose of an object. It uses a convolutional neural network (CNN) to detect the 2D bounding box of the object in the image and then applies a perspective-n-point (PnP) algorithm to estimate the 3D pose of the object based on the 2D bounding box and known geometric priors. Figure 30 shows the prediction of the 4 predefined objects.

Recent efforts in the area of 6DoF object pose estimation using RGB input have achieved state-of-the-art accuracy. However, the run-time increases linearly with the number of objects due to the need to compute the 6D posture using PnP for each object separately. Additionally, other methods employ a pixel-wise RANSAC-based voting system to find the required key points, and a separate 2D object detector may be required to locate and clip the bounding boxes of the items of interest. These factors make these approaches unsuitable for use cases involving many objects and runtime restrictions, preventing their use in a variety of real-world circumstances. Lately, generalisable pose estimators achieved promising results. These models learn to extract relevant features from images or point cloud data and predict the object's pose. Gen6D, [8], is a generalizable model-free 6D object pose estimator. Existing generalizable pose estimators either need the high-quality object 3D model or require additional depth maps or object masks in test time, which significantly limits their application scope. In contrast, this pose estimator only requires some posed images of the unseen object and can accurately predict the poses of the object in arbitrary environments.

Immersive technologies like augmented reality (AR) and virtual reality (VR) have gained a lot of traction in the healthcare field. They offer tremendous opportunities to enhance medical education, therapeutic treatments, and the patient's emotional and physical care [9], [10]. By increasing the accuracy of surgical techniques, AR and VR provide new chances to understand diseases [11], [12]. Neurosurgery is included in the scientific research on augmented reality, in addition to studies on brain tumours [13]. The introduction of a 3D interface (HI) prototype for visualising MRI data and scheduling neurosurgical procedures is a noteworthy achievement in this regard [13]. Similarly, researchers suggested employing a 3D-printed head in an interactive AR environment for neurosurgery teaching 3, the HoloLens simulator showed remarkable results after including this interface. Another study concentrated on the integration of distinct AR technology to aid navigational biopsy operations utilising a head-mounted optical display (HMD) on a personalised brain phantom [14]. Researchers looked into the viability and precision of 3D Neural Navigation using the HoloLens smart augmented reality glasses in [15]. This paper investigates how the 6 DoF head pose estimation deep learning-based method deployed into augmented reality may transform medical practice training for brain tumour diagnosis. figure 1 depicts the global flowchart of the proposed framework. Our AR platform termed 'HOLOTumour' provides AR visualisation of brain tumour segmentation results for monitoring the patient's vital signs and health conditions and providing information on the segmented brain tumour to make healthcare decisions. The principal contributions of the conceived AR-medical practice platform HOLOTumour are enumerated below:

- An improved method based on Geodesic-Aided Chan-Vese Model for brain segmentation using MRI scans.
- A 3D head model reconstruction module based on the results of brain tumour segmentation, further to 3D phantom head print.
- A Deep learning-based vision module to automatically estimate the head 6 DoF pose using only RGB images.
- HOLOTumour AR platform for medical practice training including the 3D models' visualisation, voice and gestural interaction and manipulation.

The paper is organised as follows. Section I studies the 6 DoF object pose estimation and how augmented reality can transform healthcare medical practice and training; it studies the related works to 6 DoF Neural Network-based pose estimation methods, and AR applied to brain tumour diagnosis. Section II delineates each one of the paper's contributions: the 6 DoF head pose estimation module, the brain segmentation further to the 3D model reconstruction and print. The HOLOTmour AR platform functionalities are presented in section II with the AR rendering visualisation, voice and hand gestural interaction. Meanwhile, the 6 DoF head pose estimation results further the brain tumour segmentation results along with comparison are presented in section III. Section III also introduced Augmented reality visualisation and interactive rendering. In section IV, we end by underlining the outcome achieved.

II. Our proposal: HOLOTumour- 6D head pose estimation for AR brain tumour Localisation

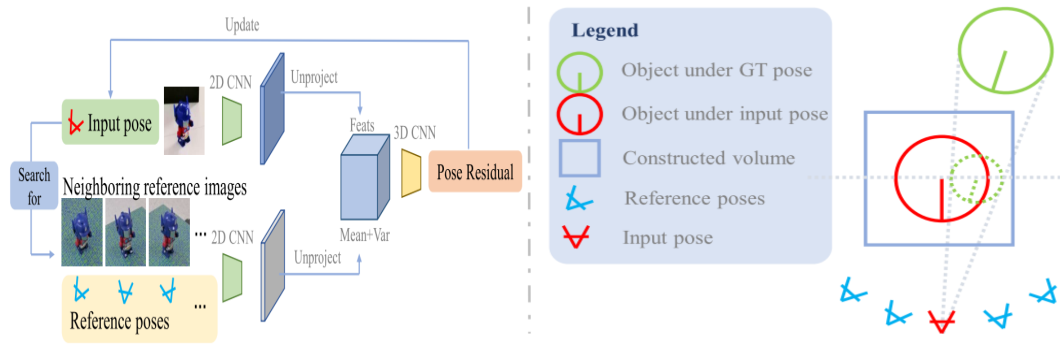
A. 6DoF pose estimation based Gen6D model

As it permits a wide range of applications inside Augmented reality (AR), being able to accurately predict the 6 DoF object pose, i.e., its 3D rotation and translation, is a key problem in computer vision. Gen6D, [8], consists of an object detector, a viewpoint selector and a pose refiner, all of which do not require the 3D object model, and can generalise to unseen objects (figure 2).

B. Geodesic-Aided Chan-Vese Model for Brain Tumor Segmentation

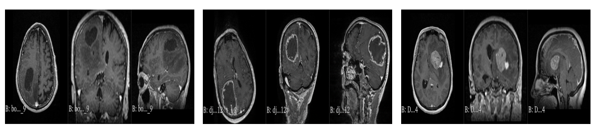
1) Dataset collection: The reputable radiology unit at Mustapha Pacha University Hospital Centre in Algeria graciously provided us with a collection of exquisite MRI scans portraying the intricate details of a patient's head and neck. These captivating scans were meticulously handpicked, adhering to stringent criteria, ensuring the presence of a magnificently proportioned and strikingly contrasting tumour. Furthermore, the scans encompassed a comprehensive MRI examination, capturing the entirety of the patient's head and neck, rendering them suitable for the enchanting realm of 3D printing. Employing a state-of-the-art Siemens Magnetom 3T Skyra scanner, the acquisition of these scans involved T1 MP-RAGE sequences, acquired post-contrast administration. The dataset comprised a grand total of 318 meticulously crafted images, arranged in axial and coronal orientations, accompanied by an additional 108 images that gracefully unfolded in the sagittal orientation. Each image, sculpted with a masterful touch, boasted a slice thickness of a mere 0.7 mm, flawlessly capturing the essence of this captivating radiological journey. Please refer to figure 3 for more details.

2) Brain tumour segmentation: The detection of brain tumours was carried out using a highly regarded segmentation technique that is considered a reference in the field of medical imaging for identifying tumour areas. In this article, the original contribution lies in a tumour segmentation method based on geodesic active contour models. These models, independently proposed by Chan and Vese [16], provide a solution that addresses the criterion of segmenting homogeneous



(a) (b)

Fig. 2: Gen6D Architecture. (a) Architecture of the pose refiner, (b) Similarity approximation [8].



(a) (b) (c)

Fig. 3: The set of actual data processed. (a) Men 58 years, (b) Men 71 years, (c) Women 50 years

regions. This criterion focuses on identifying regions where the pixel intensities closely resemble the average intensity of the respective region. By leveraging this approach, the segmentation process achieves greater accuracy and consistency in delineating tumour boundaries. The model described in Equation (1) incorporates the alignment force as a driving force for the active contour. Additionally, it integrates the geodesic active contour force for regularisation, along with the minimum variance criterion originally proposed by Chan and Vese [16].

$$E\left(C, c_1, c_2\right) = E_{AR}\left(C\right) - \alpha E_{GAC}\left(C\right) - \beta E_{MV}\left(C, c_1, c_2\right) \tag{1}$$

The two constants, c_1 and c_2, represent the mean intensities inside and outside the contour C, respectively. These constants are used in the model alongside the parameters α and β, which are also constants.

- The robust alignment term is given by the functional :

$$E_{AR}\left(C\right) = \oint_C |\langle \nabla I, \overrightarrow{n} \rangle| ds \tag{2}$$

The gradient direction, represented by ∇I, is effective in estimating the orientation of the edge contour, denoted as C. The alignment term in the model attains a higher value when the normal vector of the curve, denoted as $\overrightarrow{n}$, aligns with the direction of the image gradient. Consequently, the objective is to find a curve that maximises this alignment functional.

- The geodesic active contour term is defined by the following functional:

$$E_{GAC}\left(C\right) = \oint_C g\left(C\left(s\right)\right) ds \tag{3}$$

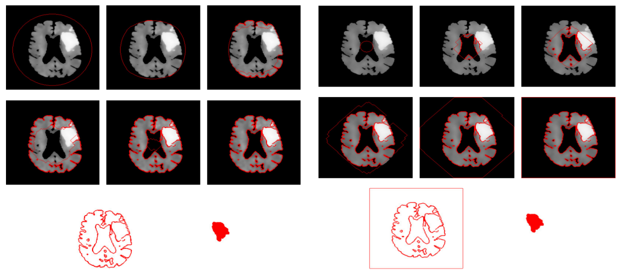
- Chen and Vese proposed a minimal variance criterion given by: [16]

$$E_{MV}\left(C, c_1, c_2\right) = \frac{1}{2}\iint_{\Omega_C} \left(I\left(x,y\right) - c_1\right)^2 dxdy + \frac{1}{2}\iint_{\Omega/\Omega_C} \left(I\left(x,y\right) - c_2\right)^2 dxdy \tag{4}$$

This function is utilised to determine the optimal separating curve. The goal is to find a curve that effectively separates the interior from the exterior regions based on their relative average values. The optimal curve is the one that achieves the best separation between these two regions.

3) Brain tumour segmentation results: To assess the effectiveness of the active geodesic contour models, we employed them on a meticulously curated dataset comprising 318-component T1 MP-RAGE sequences. The determination of the optimal number of iterations was a meticulous process, meticulously considering the size of the mass as evaluated by a seasoned medical specialist, as we observed its influential role in the iterative convergence. Within the confines of this study, the implemented segmentation method unfolded its prowess, identifying the principal brain tissues with utmost precision, while simultaneously extracting the elusive realm of tumour territory. This extraction paved the way for subsequent calculations, delving into the intricate characteristics of the tumour for the purpose of awe-inspiring 3D reconstruction and captivating visualisation. The innovative approach was meticulously crafted within the enchanting confines of MATLAB R2020b, harmonising its brilliance with the formidable prowess of an MSI computer graced by the presence of an Intel i7-9750H CPU, delicately dancing at 2.60 GHz, and accompanied by the exquisite abundance of 32 GB of RAM. The extraordinary capabilities of this proposed system, which encompass segmentation, 3D reconstruction, and visualization are gracefully showcased within the intricate fabric of the experimental findings, eagerly anticipating exploration in the upcoming subsections. The culmination of these efforts will be immortalized in the ethereal realm of a magnificent Figure 4 (a) and 4 (b), exhibiting the tangible results of our hard work and shedding light on the path to uncovering novel insights.

By testing the topology change criterion on real MRI



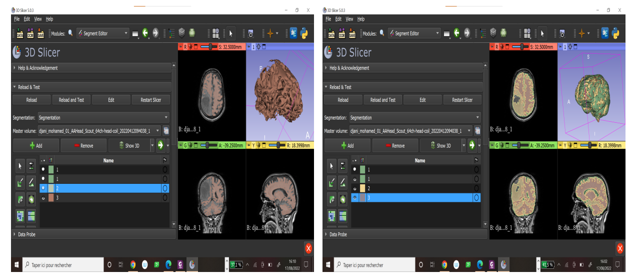
(a) (b)

Fig. 4: Process Convergence and Results of Segmentation Geodesic Active Contour Models: (a) For internal Convergence; (b) For Outward Convergence.

images, figures 4 (a) and 4 (b) demonstrate the convergence results achieved after a specific number of iterations for outward and inward segmentation. Adhering to this criterion ensures accurate segmentation outcomes while preserving the topology of the objects of interest. Increasing the variance value enables the method to capture well-defined contours of the segmented objects, considering topological changes and contour closure. This proves crucial in accurately delineating the boundaries of the objects, even when the initial active contour is initialised far from the desired objective. Moreover, the significance of proper initialisation is emphasised, as it directly impacts the segmentation process. By carefully initialising the initial active contour, the methodology overcomes initialisation biases and achieves successful segmentation outcomes. In summary, this brief discussion highlights the effectiveness of the topology change criterion and the incorporation of texture considerations in MRI image segmentation using the Chan and Vese method. The presented results contribute to advancing segmentation techniques in medical imaging, enhancing the accuracy and reliability of object delineation in MRI analysis.

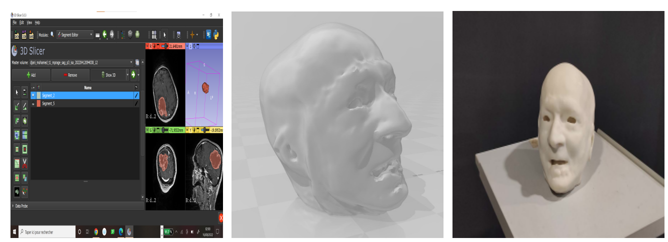
4) 3D reconstruction: Utilising the cutting-edge Slicer module [17], ”Tsallis Brain Segmentation,” we delve into the realm of brain reconstruction. This remarkable tool extracts and segments the brain, specifically targeting the elusive grey matter. However, as revealed in figure 5 (a), its limitations arise when confronted with the intricate distortions caused by the presence of tumours. To surmount this obstacle, we turn to the ingenious ”Segment Editor” module, employing its ”wrap solidify” segmentation technique to fill in the gaps. Behold, as displayed in figure 5 (b), the tumour-ridden regions of the brain are now seamlessly filled and fortified.

5) Phantom head print: In order to test and evaluate our application, a layering procedure was developed by working on a dummy head fabricated using an sPro 60 3D printer, which uses a 3D printing technology called selective laser sintering (SLS). We examined a case containing a brain tumor of a patient obtained from a volunteer from the University Hospital of Bab El-Oued, Algeria, by means of anatomical MRI as shown in figure (a) 6. Specifically, a virtual 3D model of the head was created using the open source 3D Slicer anatomy software see figure (b) 6, in order to convert the MRI images of the brain and save them to a .stl file (Standard Tessellation Language) as shown in figure (c) 6. This file format is widely used for rapid prototyping, 3D printing, and computer-aided design (CAD). STL files only describe the surface geometry of a 3D object, with no representation of color, texture, or other attributes common to CAD models. The material used for printing the head is Nylon PA12 (plastic), which prints objects from polyamide powder. Among the properties of this material are strength and flexibility, as well as resistance to stress when bending. It also allows complex forms to be printed cheaply and quickly. We can see the final result of the printed model in Figure 4.



(a) (b)

Fig. 5: Brain reconstruction: (a) Tsallis brain segmentation for brain extraction; (b) wrap solidifyr for brain unification.



(a) (b) (c)

Fig. 6: Stages of creating a virtual 3D model of the head: (a) Insert MRI images into a 3D Slicer after segmentation of the tumor with Chan and Vese method; (b) Ready STL file for 3D printing; (c) Phantom head following print.

C. HOLOTumour AR platform

1) Gen6D deployment for Pose estimation augmented rendering: Our preliminary visual results for phantom head - DoF pose estimation based Gen6D are displayed in figure 7. The rendered images of the unseen object ’phantom head show the 6 DoF pose estimation from different viewpoints.

2) Virtual Model Overlay: We used Blender [18] to treat the phantom head 3D model, fixing any errors that may occur on extraction. Unity, [**?**], gives complete access to any object created and imports 3D modules (.obj), which we need to import our brain and segmented tumour 3D modules in our AR scene. Figure 8 displays the visualisation results of the phantom head virtual overlay.

III. HOLOTUMOUR RESULTS

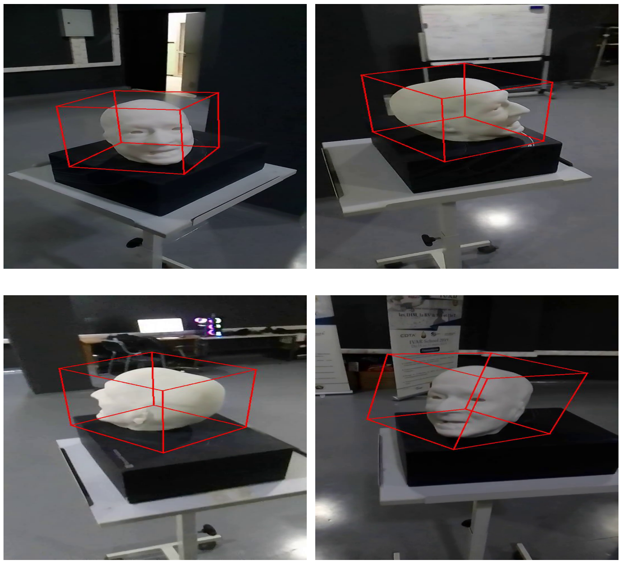


Fig. 7: Gen6D performance on our phantom head.

Fig. 8: Phan head virtual overlay

A. 6D pose estimation based DL results and comparison

In this part, we compare four deep learning-based 6 DoF object pose estimators: EfficientPose [4], PvNet [3], Gen6D [8], and OnePose++ [2]. We consider the widely-used Average Distance (ADD) [19] (ADD equation 5, and ADD(-S) equation 6) and the projection error as metrics. We compute the recall rate on the ADD with 10% of the object diameter (ADD-0.1d). On the projection error, we calculate the recall rate at 5 pixels (Prj-5). We used a workstation with Intel(R) Core(TM) i9-10900F CPU 2.81GHz, 32 Go of RAM and NVIDIA GeForce RTX 3070 Ti graphical card.

$$ADD = \frac{1}{m}\sum_{x \epsilon M}\left\|(Rx+t)-(\tilde{R}x+\tilde{t})\right\|_2 \quad (5)$$

TABLE I: Gen6D performance on LINEMOD dataset [7]

Dataset	Metrics	Cat	Phone	Lamp	Iron	Holepuncher	Glue
LINEMOD dataset	ADD-0.1d	0.6008	0.8156	0.8964	0.9397	0.6765	0.5232
[7]	Proj-5	0.9611	0.9664	0.9155	0.9540	0.9867	0.9633

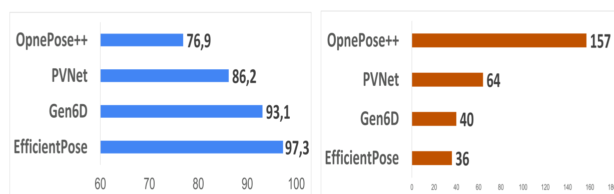
TABLE II: Gen6D performance on GenMop dataset [8]

Dataset	Metrics	Chair	TFormer	Miffy	Head	Marqueur	Cup
GenMop dataset	ADD-0.1d	0.6150	0.4762	0.6111	0.0303	0.4286	0.7622
[8]	Proj-5	0.5500	0.9643	1.0000	0.0061	0.1293	0.8649

$$ADD-S = \frac{1}{m}\sum_{x_1 \epsilon M} min_{x_2 \epsilon M}\left\|(Rx+t)-(\tilde{R}x+\tilde{t})\right\|_2 \quad (6)$$

This metric calculates the average point distances between the 3D model point set M transformed with the ground truth rotation R and translation t and the model point set transformed with the estimated rotation $\tilde{R}$ and translation $\tilde{t}$.

Figure 9-(a) is a graph comparing the ADD(-S) accuracy scores of the methods on the LINEMOD dataset [7], which is the most common dataset used for 3D object pose estimation evaluations. Figure 9-(b) is a graph comparing the running time (ms) of the four methods during the inference phase, which is the most important criterion, because the lower the runtime the better the FPS. All these runtimes were calculated while running inference using a GPU, which produces 10-25 FPS, running inference on a CPU will result in much slower runtimes (around 2-5 FPS), which is due to the neural network's complexity. This limits their use in PCs without a dedicated GPU and most mobile devices.



(a) (b)

Fig. 9: Comparative performance LINEMOD dataset [7] . (a) Comparative of Accuracy Scores ADD(-S), (b) Comparative of Inference Runtimes.

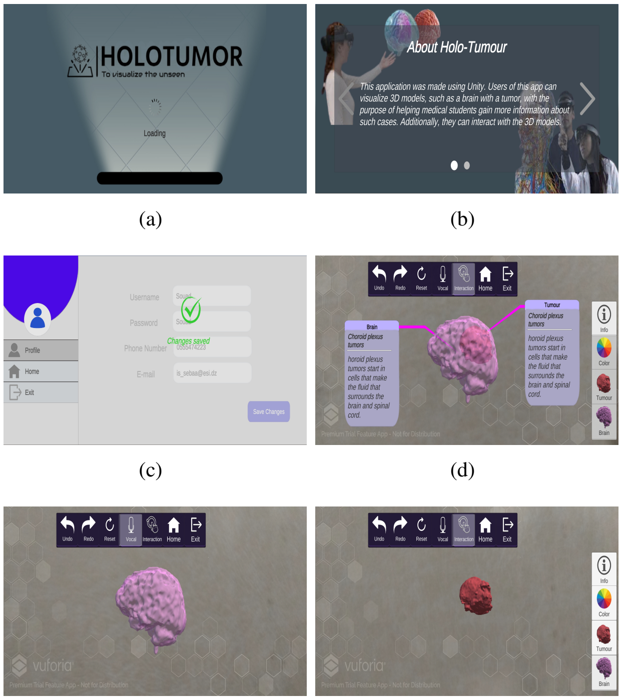
Experiments show that Gen6D achieves state-of-the-art results on two model-free datasets: the MOPED dataset [20] (table I and a new GenMOP dataset [8] (table II. In addition, on the LINEMOD dataset [7], Gen6D achieves competitive results compared with instance-specific pose estimators.

B. AR visualisation and interaction results

Figure 10 showcases the HOLOTumour functionalities including user account creation, deletion and modification, AR brain visualisation, and AR brain tumour visualisation. Figure 11 depicts the voice and gestural interaction visual results.

IV. Conclusion

The 6 DoF object pose estimation for AR applications is challenging due to the variety of objects and the complexity of a scene caused by clutter and real-time task requirement.



(a) (b) (c) (d) (e) (f)

Fig. 10: HOLOTumour results. (a) HOLOTumour logo, (b) home page, (c) Profile changes, (d) brain and segmented tumour visualisation, (e) brain visualisation using the vocal command, (f) brain tumour visualisation using hand command

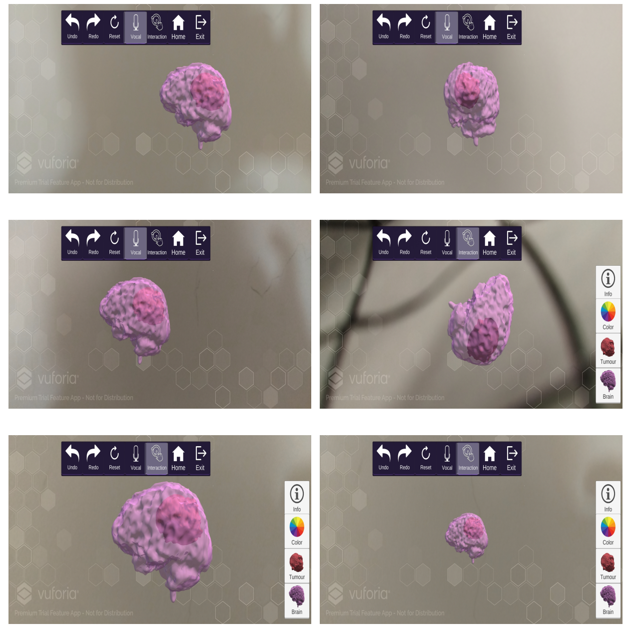


Fig. 11: HOLOTumour AR gestural and vocal interaction results.

In this work, we introduced HOLOTumour, an AR platform that segments brain tumours using Geodesic-Aided Chan-Vese Model, estimates the 6 DoF phantom printed head pose, and provides an AR rendering for visualisation and interaction using voice and gestural command. For the 6 DoF phantom head estimation, we choose the Gen6D estimator. It is Model-free that only takes RGB images, and generalisable and can be applied on an arbitrary object without training on the object or its category. As a result and based on the comparison, Gen6D was able to estimate the 6 DoF phantom head pose and brain tumour segmentation results were satisfactory. Our results are quite encouraging in deploying Gen6D model into AR visualisation and interaction for realistic rendering. The HOLOTumour platform, which may be used for professional practice and training in medicine, can assist medical practitioners in identifying brain tumours and visualising the augmented segmented brain tumour rendering.

REFERENCES

[1] A. Collet, M. Martinez, and S. Srinivasa, "The moped framework: Object recognition and pose estimation for manipulation," *I. J. Robotic Res.*, vol. 30, pp. 1284–1306, 09 2011.

[2] X. He, J. Sun, Y. Wang, D. Huang, H. Bao, and X. Zhou, "Onepose++: Keypoint-free one-shot object pose estimation without cad models," 01 2023.

[3] S. Peng, Y. Liu, Q. Huang, X. Zhou, and H. Bao, "Pvnet: Pixel-wise voting network for 6dof pose estimation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[4] Y. Bukschat and M. Vetter, "Efficientpose: An efficient, accurate and scalable end-to-end 6d multi object pose estimation approach," *CoRR*, vol. abs/2011.04307, 2020. [Online]. Available: https://arxiv.org/abs/2011.04307

[5] M. L. Solutions, "Objectron (3d object detection)," 2023. [Online]. Available: https://github.com/google/mediapipe/blob/master/docs/solutions/objectron.md

[6] X. He, J. Sun, Y. Wang, D. Huang, H. Bao, and X. Zhou, "Onepose++: Keypoint-free one-shot object pose estimation without cad models," in *Advances in Neural Information Processing Systems*, S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, Eds., vol. 35. Curran Associates, Inc., 2022, pp. 35 103–35 115. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2022/file/e43f900f571de6c96a70d5724a0fb565-Paper-Conference.pdf

[7] S. S. G. K. S.Hinterstoisser, V.Lepetit and N.Navab, "Model-basedtraining,detection and pose estimation of texture-less 3d objects in heavily cluttered scenes," in *Computer Vision–ACCV2012*, Springer, Ed., 2013, p. 548–562.

[8] Y. Liu, Y. Wen, S. Peng, C. Lin, X. Long, T. Komura, and W. Wang, "Gen6d: Generalizable Model-Free 6-DoF Object Pose Estimation from RGB Images," 2022. [Online]. Available: https://arxiv.org/abs/2204.10776

[9] K. Amara, N. Ramzan, N. Achour, M. Belhocine, C. Larbas, and N. Zenati, "Emotion recognition via facial expressions," in *2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA)*, 2018, pp. 1–6.

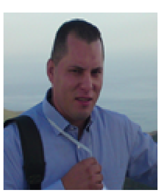
[10] K. Amara, N. Ramzan, N. Zenati, O. Djekoune, C. Larbes, M. Guerroudji, and D. Aouam, "Towards emotion recognition in immersive virtual environments: a method for facial emotion recognition," in *ICCSA 2021 Conference on Computer Science's Complex Systems and their Applications 2021*, ser. CEUR Workshop Proceedings, T. Marir, A. Bourouis, R. Benaboud, V. Gupta, and C. Gupta, Eds., vol. 2904. Germany: CEUR Workshop Proceedings, May 2021, pp. 253–263, 2nd International Conference on Computer Science's Complex Systems and their Applications, ICCSA 2021 ; Conference date: 25-05-2021 Through 26-05-2021.

[11] K. Amara, A. Aouf, H. Kennouche, A. O. Djekoune, N. Zenati, O. Kerdjidj, and F. Ferguene, "Covir: A virtual rendering of a novel nn architecture o-net for covid-19 ct-scan automatic lung lesions segmentation," *Computers Graphics*, vol. 104, pp. 11–23, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0097849322000358

[12] K. Amara, H. Kennouche, A. Aouf, O. Kerdjidj, N. Zenati, O. Djekoune, and M. A. Guerroudji, "Augmented reality for medical practice: a

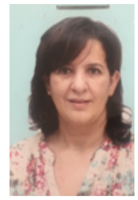
comparative study of deep learning models for ct-scan segmentation," in *2022 International Conference of Advanced Technology in Electronic and Electrical Engineering (ICATEEE)*, 2022, pp. 1–6.
[13] C. M. Morales Mojica, N. V. Navkar, N. V. Tsekos, D. Tsagkaris, A. Webb, T. Birbilis, and I. Seimenis, "Holographic interface for three-dimensional visualization of mri on hololens: A prototype platform for mri guided neurosurgeries," in *2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE)*, 2017, pp. 21–27.
[14] M. Majak, M. Żuk, E. Swiatek-Najwer, M. Popek, and P. Pietruski, "Augmented reality visualization for aiding biopsy procedure according to computed tomography based virtual plan," *Acta of bioengineering and biomechanics / Wroclaw University of Technology*, vol. 23, p. 2021, 03 2021.
[15] T. Doormaal, J. van Doormaal, and T. Mensink, "Clinical accuracy of holographic navigation using point-based registration on augmented-reality glasses," *Operative Neurosurgery*, vol. 17, 05 2019.
[16] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on image processing*, vol. 10, no. 2, pp. 266–277, 2001.
[17] "3d slicer," *3D Slicer*, 2023. [Online]. Available: https://www.slicer.org
[18] "Blender," *Blender*, 2023. [Online]. Available: https://www.blender.org
[19] S. Hinterstoisser, V. Lepetit, S. Ilic, S. Holzer, G. Bradski, K. Konolige, and N. Navab, "Model based training, detection and pose estimation of texture-less 3d objects in heavily cluttered scenes," vol. 7724, 10 2012.
[20] K. Park, A. Mousavian, Y. Xiang, and D. Fox, "Latentfusion: End-to-end differentiable reconstruction and rendering for unseen object pose estimation," 2020.

Amara. K obtained her PhD degree at University of science and technology Houari Boumediene USTHB, Algiers, Algeria, in 2018 and her master in control and robotics at USTHB in 2011. She works as a permanent researcher at CDTA with the IRVA team. Her main research includes: Augmented and Virtual Reality, 3D Interaction, Affective Computing, Computer Vision, Emotion Recognition, Healthcare.



Guerroudji. M.A Mohamed Amine Guerroudji obtained PhD degree at UMMTO, Tizi ouzou, Algeria, in 2017 and master in information processing and computer system at USDB, Blida, Algeria, in 2013. He works as a permanent researcher at CDTA with the IRVA team. main research includes augmented and virtual reality, 3D interaction, and computer vision.

Kerdjidj. O obtained his PhD degree at the University of Laghoaut, Algeria, in 2019. His leading research includes Hardware and Software implementation and Artificial Intelligence applied to Healthcare applications.



Zenati. N received her PhD degree in Process Control and Robotics from the University of Franche-Comté, France, in 2008. Now, she works as a research director at CDTA. Her main research includes Augmented and Virtual Reality, 3D Interaction, e-maintenance, and Healthcare.

Ramzan. N is a Full Professor of Computing Engineering and Chair of Affective and Human Computing for Smart Environment Research Centre and Co-lead Visual Communication Cluster in AVCN at UWS. He is currently focused on leading high-quality interdisciplinary research and teaching in the area of video processing, analysis and communication, multimedia search and retrieval, video quality evaluation, brain-inspired multi-modal cognitive technology, multi-modal human-computer interfaces, DNA computing, fall detection, big data analytics, affective computing, IoT/smart environments, natural multi-modal human-computer interaction, and eHealth/connected Health. He received his M.Sc. degree in telecommunication from University of Brest, France, in 2004 and his PhD degree in electronics engineering from Queen Mary University of London, London, U.K., in 2008. Before joining UWS, he was a senior research fellow and lecturer at Queen Mary University of London from 2008 to 2012. He is, a senior member of the IEEE, Fellow of the Higher Education Academy (HEA), co-chair of MPEG HEVC verification (AHG5) group and a voting member of the British Standard Institution (BSI). In addition, he holds key roles in the Video Quality Expert Group (VQEG) such as Co-chair of the Ultra High Definition (UltraHD) group; Co-chair of the Visually Lossless Quality Analysis (VLQA) group; and Co-chair of the Psycho-Physiological Quality Assessment (PsyPhyQA). He has published more than 90 articles in journals, conferences, and book chapters including standard contributions. He has been awarded a STARS (Staff Appreciation and Recognition Scheme) award for 2014 and 2016 for "Outstanding Research and Knowledge Exchange" (University of the West of Scotland) and awarded Contribution Reward Scheme 2011 and 2009 for "Outstanding Research and Teaching Activities" (Queen Mary University of London).